

TWO WEIGHT L^p ESTIMATES FOR PARAPRODUCTS IN NON-HOMOGENEOUS SETTINGS

JINGGUO LAI AND SERGEI TREIL

ABSTRACT. We give a necessary and sufficient condition for the two weight L^p -estimates for paraproducts in non-homogeneous settings, $1 < p < \infty$. We are mainly interested in the case $p \neq 2$, since the case $p = 2$ is a well-known and easy corollary of the Carleson embedding theorem. The necessary and sufficient condition is given in terms of testing conditions of Sawyer type: for $p \leq 2$ only one (“direct”) testing condition is required, but for $p > 2$ both “direct” and “adjoint” testing conditions are needed. An interesting feature is that the “adjoint” testing condition is that it is a testing condition not for the adjoint of the paraproduct, but for some auxiliary operator.

CONTENTS

1. Introduction and main results	1
2. The case $1 < p \leq q$: the sufficiency	7
3. The case $2 < p < \infty$: a counterexample	9
4. The case $2 < p < \infty$: reduction to the shifted positive martingale operators and necessity of condition (1.14)	11
5. Two weight estimates for shifted positive martingale operators	15
References	22

1. INTRODUCTION AND MAIN RESULTS

The paper is devoted to the two weight estimates of paraproducts in the non-homogeneous martingale situation. Paraproducts play an important role in the investigation of the weighted inequalities for the singular integral operators. The L^2 -boundedness of paraproducts is easy, a necessary and sufficient condition follows immediately from the Carleson embedding theorem.

This necessary and sufficient condition can be stated as a *testing condition*, i.e. a dyadic paraproduct is bounded in L^2 if and only if there is a uniform estimate on all dyadic intervals (cubes). In the classical homogeneous (non-weighted) situation the L^2 boundedness (or L^{p_0} boundedness, for some $1 < p_0 < \infty$) is equivalent to the boundedness of the paraproduct in *all* L^p , $1 < p < \infty$.

The non-homogeneous situation is much more interesting, even in the non-weighted case. Namely, it was shown in [19] that in this case the testing condition is still necessary and

2010 *Mathematics Subject Classification.* 42B20, 42B25, 60G42.

Key words and phrases. paraproducts, non-homogeneous space, two weight estimates.

Supported in part by the National Science Foundation under the grant DMS-1301579.

sufficient, but it now depends on p : the boundedness in L^{p_0} implies the boundedness in L^p with $1 < p \leq p_0$ but not in L^p with $p_0 < p < \infty$.

Two weight case becomes even more interesting: while it is not hard to show that for $p \leq 2$ the testing condition is still sufficient for the boundedness, we will present a counterexample showing that for $p > 2$ the testing condition alone does not work.

It will be shown in the paper that adding a second, “dual” testing condition we get necessary and sufficient conditions for the two weight estimates for $p > 2$.

It has been understood for a long time, that in many cases one needs two testing conditions, one for the operator, the other one for the adjoint, to get a sufficient condition for the boundedness. That was the case, for example for many positive operators (i.e. integral operators with positive kernels), in particular for the so-called *positive dyadic operators*. That was also the case for Calderón–Zygmund operators, but in this case the bounds on paraproduct followed from one testing condition, and the dual testing condition implied bounds on the adjoint paraproduct.

But in our case we need 2 testing condition for the estimates of paraproduct! Also, our “dual” testing condition is not a testing condition for the adjoint paraproduct, but for some auxiliary operator.

1.1. Preliminaries. The theory of L^2 two weight estimates is now well developed for many classes of singular operators, including the so-called “well localized” operators, treated in [12] where necessary and sufficient condition of Sawyer type was proved for such operators.

The Sawyer type conditions are essentially the testing conditions on the operator and its adjoint, meaning that to prove that the operator is bounded, it is sufficient to check that the operator and its adjoint are (uniformly) bounded on characteristic functions $\mathbf{1}_Q$ of cubes Q . We say “essentially” here, because in many cases this testing condition can be weakened a bit, like restricting the result only to the cube Q , or in case of martingale operators also considering smaller sums.

The recent achievements in the theory of L^2 two weight estimates include sufficiency of Sawyer type estimates for Hilbert Transform [7], [5], for the Cauchy Transform [8], for the Riesz transforms [9].

As for the L^p theory for $p \neq 2$ everything is well understood for the “positive” operators, including the maximal function [14], fractional and Poisson integrals [15], the so-called positive dyadic operators [6], [18].

As for the more “singular” operators we should first mention [13], where a necessary and sufficient condition for the two weight estimates of the vector-valued martingale operators was obtained. Such operators can be interpreted as a simplified “model” for paraproducts, and the necessary and sufficient conditions were “kind of” Sawyer type conditions. More precisely, the “direct” condition was exactly the Sawyer condition for the operator, but the adjoint condition was a more restrictive one, in a sense that for each interval I one has to check the boundedness not on one function, but on an infinite family of (vector-valued) functions.

Very interesting recent papers [20], [21] give necessary and sufficient conditions for the two weight estimates of the well localized operators and of the Haar shifts, but again the conditions were more restrictive than the Sawyer type ones: again, for each interval I one has to check the boundedness on infinitely many vector-valued functions.

In this paper we consider paraproducts and the corresponding vector-valued martingale operators and get necessary and sufficient condition for the L^p two weight estimates. Our conditions are exactly Sawyer type conditions, meaning that for each interval we have to test only one function $\mathbf{1}_I$. However, while the “direct” condition is exactly the testing condition for the original operator, the “adjoint” one tests an auxiliary operator, acting not between duals of the original L^p spaces, but between the duals of $L^{p/2}$ spaces. We should mention that for $p \leq 2$ the “direct” testing conditions alone is sufficient; for $p > 2$ both “direct” and “adjoint” conditions are required.

1.2. The setup. The general setup is as follows. We consider a σ -finite measure space $(\mathcal{X}, \mathfrak{S}, \mu)$ with a filtration (i.e. with increasing sequences of σ -algebras) $\mathfrak{S}_n \subset \mathfrak{S}_{n+1} \subset \mathfrak{S}$, $n \in \mathbb{Z}$.

We assume that each σ -algebra \mathfrak{S}_n is *atomic*, i.e. that there exist a countable collection \mathcal{L}_n of disjoint sets of finite measure (atoms) such that each $A \in \mathfrak{S}_n$ is a union of sets $I \in \mathcal{L}_n$.

The fact that $\mathcal{L}_n \subset \mathcal{L}_{n+1}$ means that each $I \in \mathcal{L}_n$ is at most countable union of $I' \in \mathcal{L}_{n+1}$.

We denote by $\mathcal{L} = \bigcup_{n \in \mathbb{Z}} \mathcal{L}_n$ the collection of all atoms (in all generations). We will also use the word *intervals* for the atoms $I \in \mathcal{L}$.

The standard example will be the filtration given by a dyadic lattice in \mathbb{R}^d but with the underlying measure being an arbitrary Radon measure σ .

We will allow a situation when an atom I belongs to several (even infinitely many) generations \mathcal{L}_n . However, we will not allow I to be in all generations, because in this case nothing interesting happens on the interval I .

There are two (equivalent) ways to define paraproduct. One is a “probabilistic” approach, when one uses discrete time n (index in \mathfrak{S}_n) and considers standard in probability martingale differences; in this approach by an atom $I \in \mathcal{L}_n$ we understand the pair (I, n) . The other one is a “geometric” approach, where one does not use the time n , and treats atoms I just as sets.

Let us start with the “probabilistic” approach. For $I \in \mathcal{L}_n$ define its *martingale* or *time* children by

$$\text{ch}^t(I, n) := \{I' \in \mathcal{L}_{n+1} : I' \subset I\}$$

For an interval $I \in \mathcal{L}$ and a measure ν such that $\nu(I) < \infty$ define the averaging operator \mathbb{E}_I^ν

$$(1.1) \quad \mathbb{E}_I^\nu f = \langle f \rangle_{I, \nu} \mathbf{1}_I = \left(\nu(I)^{-1} \int_I f d\nu \right) \mathbf{1}_I;$$

we use the notation $\langle f \rangle_{I, \nu}$ for the average

$$\langle f \rangle_{I, \nu} := \nu(I)^{-1} \int_I f d\nu$$

with the understanding that $\langle f \rangle_{I, \nu} = 0$ if $\nu(I) = 0$.

We define the martingale difference $\Delta_{(I, n)}^\nu$ as

$$\Delta_{(I, n)}^\nu := \sum_{I' \in \text{ch}(I, n)} \mathbb{E}_{I'}^\nu - \mathbb{E}_I^\nu$$

Let μ, ν be locally finite measures (i.e. finite on all $I \in \mathcal{L}$) and let b be ν -integrable on each $I \in \mathcal{L}$. Define the two weight paraproduct $\pi = \pi_b = \pi_b^{\mu, \nu}$ by

$$(1.2) \quad \pi_b^{\mu, \nu} f = \sum_{n \in \mathbb{Z}} \sum_{I \in \mathcal{L}_n} \left(\mathbb{E}_I^\mu f \right) \left(\Delta_{(I, n)}^\nu b \right).$$

Note that if an interval I belongs to several \mathcal{L}_n , all the term in the paraproduct corresponding to this interval, with a probable exception of the term with the largest n are 0. So we can write the paraproduct differently, not writing zero terms.

Namely, one can define the rank $\text{rk } I$ of $I \in \mathcal{L}$ as supremum of n such that $I \in \mathcal{L}_n$ (it can be equal $+\infty$). Then for $I \in \mathcal{L}$, $\text{rk}(I) = n$ define the children $\text{ch}(I)$ to be the collection of atoms $I' \in \mathcal{L}_{n+1}$ such that $I' \subset I$; if $\text{rk}(I) = \infty$ we set $\text{ch}(I) := \{I\}$.

Then we can define the martingale difference Δ_I as

$$\Delta_I^\nu := \sum_{I' \in \text{ch}(I)} \mathbb{E}_{I'}^\nu - \mathbb{E}_I^\nu$$

and rewrite the paraproduct $\pi = \pi_b = \pi_b^{\mu, \nu}$ as

$$(1.3) \quad \pi_b^{\mu, \nu} f = \sum_{n \in \mathbb{Z}} \sum_{I \in \mathcal{L}_n} \left(\mathbb{E}_I^\mu f \right) \left(\Delta_I^\nu b \right).$$

This is what we call the *geometric* notation for paraproducts.

Remark 1.1. All results and proofs in the paper are valid in both “geometric” and “probabilistic” settings. For the sake of brevity of writing we will use the “geometric” notation. However, we should emphasize that everything can be easily (formally) translated to the “probabilistic” setup: one just has to understand an atom $I \in \mathcal{L}_n$ as a pair (I, n) , and inclusion $I \subset J$ for atoms $I \in \mathcal{L}_n$ and $J \in \mathcal{L}_k$ should be understood as inclusion for sets plus the inequality $n \geq k$. The union or intersection of atoms in this case should be understood as union/intersection of the corresponding sets and the “time” n should not be taken into account.

Remark 1.2. When we introduced paraproduct we assumed that we are given a locally integrable function b (symbol of the paraproduct) whose martingale differences $\Delta_I^\nu b$ are involved in the definition. However, that is not necessary, the results of the paper hold if we just assume that we are given a sequence of martingale differences $b_I = \Delta_I^\nu b$, without assuming that they correspond to a function b .

We only assume that each martingale difference b_I is an integrable function supported on I , constant on children of I and ν -orthogonal to constants, i.e. that $\int_I b_I d\nu = 0$.

In this case we will still use b for the symbol of the paraproduct, but we understand b as the collection $b = \{b_I\}_{I \in \mathcal{L}}$ of all martingale differences. But we will still use the symbol $\Delta_I^\nu b$ for b_I to remind us about the origins.

1.3. Main results. In this paper we investigate when $\pi_b^{\mu, \nu}$ is a bounded map from $L^p(\mu)$ to $L^p(\nu)$ for $1 < p < \infty$.

1.3.1. Estimates for paraproducts. For $I \in \mathcal{L}$ denote $\mathcal{L}(I) := \{I' \in \mathcal{L} : I' \subset I\}$.

Theorem 1.3. *Let $1 < p \leq 2$. Then the estimate*

$$(1.4) \quad \|\pi_b^{\mu, \nu} f\|_{L^p(\nu)}^p \leq A^p \|f\|_{L^p(\mu)}^p,$$

holds if and only if for all $J \in \mathcal{L}$

$$(1.5) \quad \int_J \left| \sum_{I \in \mathcal{L}(J)} \Delta_I^\nu b \right|^p d\nu \leq B^p \mu(J).$$

Moreover, for the best constants we have $c(p)B \leq A \leq C(p)B$, where the constants $c(p)$, $C(p)$ depends only on p .

Theorem 1.4. *Let $2 < p < \infty$. The estimate (1.4) holds if and only if (1.5) and*

$$(1.6) \quad \int_J \left[\sum_{I \in \mathcal{L}(J)} \frac{\nu(I)}{\mu(I)} \mathbb{E}_I^\nu \left(|\Delta_I^\nu b|^2 \right) \right]^{r'} d\mu \leq B_*^{r'} \nu(J), \quad r = p/2$$

hold for all $J \in \mathcal{L}$.

Moreover, for the best constants we have $c(p) \max\{B, B_*\} \leq A \leq C(p) \max\{B, B_*\}$.

Remark 1.5. The one weight case $\nu = \mu$ was treated in [19], where it was shown that (1.4) holds if and only if (1.5) holds for all $1 < p < \infty$. However, in the two weight case, as we will see below in Section 3, the condition (1.5) alone is not sufficient if $2 < p < \infty$.

1.3.2. Littlewood–Paley inequalities and estimates of generalized vector paraproducts. For a function $f \in L^p(\nu)$, $1 \leq p \leq \infty$ define its *square function*

$$(1.7) \quad S^\nu f := \left(\sum_{I \in \mathcal{L}} |\Delta_I^\nu f|^2 \right)^{1/2}.$$

We will need the following well-known result.

Theorem 1.6 (Littlewood–Paley estimates). *For $f \in L^p(\nu)$, $1 < p < \infty$*

$$\|S^\nu f\|_{L^p(\nu)} \leq C(p) \|f\|_{L^p(\nu)}.$$

Moreover, if a function $f \in L^p(\mu)$ is decomposed $f = \sum_{I \in \mathcal{L}} \Delta_I^\nu f$, then

$$c(p) \|f\|_{L^p(\nu)} \leq \|S^\nu f\|_{L^p(\nu)}.$$

Here the constants $0 < c(p) < C(p) < \infty$ depend only on p .

This theorem allows us to reduce Theorems 1.3 and 1.4 to estimates of generalized vector-valued paraproducts.

Following [19] define the Triebel–Lizorkin type space $\dot{\mathbf{g}}_p^q(\mathcal{L}, \nu)$ as the space of sequences $s = \{s_I\}_{I \in \mathcal{L}}$ such that

$$\|s\|_{\dot{\mathbf{g}}_p^q(\mathcal{L})} := \left\| \left(\sum_{I \in \mathcal{L}} |s_I|^q \mathbf{1}_I \right)^{1/q} \right\|_{L^p} < \infty.$$

For $I \in \mathcal{L}$ let \widehat{I} denote the parent of I . Let $\beta = \{\beta_{I, I'}\}_{I \in \mathcal{L}, I' \in \text{ch}(I)}$ be a numerical sequence. Define the *generalized vector paraproduct* Π_β acting from $L^p(\mu)$ to the space of sequences,

$$\Pi_\beta f = \left\{ \langle f \rangle_{\widehat{I}, \mu} \beta_{\widehat{I}, I} \right\}_{I \in \mathcal{L}}$$

Remark 1.7. By the Littlewood–Paley estimates (Theorem 1.6) we can see that the estimate (1.4) is equivalent (up to a constant depending only on p) to the bounds on the operator $\Pi_\beta : L^p(\mu) \rightarrow \dot{\mathbf{g}}_p^2(\mathcal{L}, \nu)$, where $\beta_{\hat{I}, I}$ is the value of $\Delta_{\hat{I}}^\nu$ on I .

Theorem below hold for arbitrary operators Π_β , not only of the ones that came from the paraproducts $\pi_b^{\mu, \nu}$. For $I \in \mathcal{L}$ define

$$(1.8) \quad b_I := \sum_{I' \in \text{ch}(I)} \beta_{I, I'} \mathbf{1}_{I'}.$$

Theorem 1.8. *Let $1 < p \leq 2$. Then the estimate*

$$(1.9) \quad \|\Pi_\beta f\|_{\dot{\mathbf{g}}_p^2(\nu)}^p \leq A^p \|f\|_{L^p(\mu)}^p \quad \forall f \in L^p(\mu),$$

holds if and only if for all $J \in \mathcal{L}$

$$(1.10) \quad \int_J \left(\sum_{I \in \mathcal{L}(J)} |b_I|^2 \right)^{p/2} d\nu \leq B^p \mu(J).$$

Moreover, for the best constants we have $B \leq A \leq C(p)B$, where the constant $C(p)$ depends only on p .

Theorem 1.9. *Let $2 < p < \infty$. The estimate (1.9) holds if and only if (1.10) and*

$$(1.11) \quad \int_J \left[\sum_{I \in \mathcal{L}(J)} \frac{\nu(I)}{\mu(I)} \mathbb{E}_I^\nu (|b_I|^2) \right]^{r'} d\mu \leq B_*^{r'} \nu(J), \quad r = p/2$$

hold for all $J \in \mathcal{L}$.

Moreover, for the best constants we have $\max\{B, c(p)B_^{1/2}\} \leq A \leq C(p) \max\{B, B_*^{1/2}\}$.*

Theorems 1.8 and 1.9 are particular cases of the theorems below.

Theorem 1.10. *Let $1 < q < \infty$, $1 < p \leq q$. Then the estimate*

$$(1.12) \quad \|\Pi_\beta f\|_{\dot{\mathbf{g}}_p^q(\nu)}^p \leq A^p \|f\|_{L^p(\mu)}^p \quad \forall f \in L^p(\mu),$$

holds if and only if for all $J \in \mathcal{L}$

$$(1.13) \quad \int_J \left(\sum_{I \in \mathcal{L}(J)} |b_I|^q \right)^{p/q} d\nu \leq B^p \mu(J).$$

Moreover, for the best constants we have $B \leq A \leq C(p, q)B$, where the constant $C(p, q)$ depends only on p and q .

Remark 1.11. The above Theorem 1.10 was proved in [1] even in higher generality: the functions $b_I \geq 0$ there could be arbitrary measurable functions supported on I , and the case $q = \infty$ was also treated there.

For the convenience of the reader we present a proof here: the proof essentially the same as the proof in [1], although in [1] due to a clever choice of the stopping parameter a better constant is obtained.

Theorem 1.12. *Let $1 < q < \infty$, $q < p < \infty$. The estimate (1.12) holds if and only if (1.13) and*

$$(1.14) \quad \int_J \left[\sum_{I \in \mathcal{L}(J)} \frac{\nu(I)}{\mu(I)} \mathbb{E}_I^\nu(|b_I|^q) \right]^{r'} d\mu \leq B_*^{r'} \nu(J), \quad r = p/q, \quad 1/r + 1/r' = 1$$

hold for all $J \in \mathcal{L}$.

Moreover, for the best constants we have $\max\{B, c(p, q)B_*^{1/q}\} \leq A \leq C(p, q) \max\{B, B_*^{1/q}\}$.

1.3.3. Trivial implications and plan of the paper. Necessity of condition (1.13) and the inequality $B \leq A$ in Theorem 1.10 is easy: one just need to test (1.12) on the function $\mathbf{1}_J$. It is then easy to see that the left hand side of (1.13) is majorated by $\|\Pi_\beta \mathbf{1}_J\|_{\dot{\mathbf{g}}_p^q(\mu)}^p$ (the latter includes some extra terms in the sum).

By the Littlewood–Paley estimates Theorems 1.3 and 1.4 are equivalent Theorems 1.8 and 1.9 respectively for the special choice of β given in Remark 1.7. By the Littlewood–Paley estimates (Theorem 1.6) condition (1.5) is equivalent (up to a constant depending only on p) to condition (1.10) for this special choice of β . Condition (1.6) is exactly condition (1.11) for this special choice of β .

Necessity of condition (1.14) will be explained later in Section 4.

Proof of sufficiency in Theorem 1.10 is quite easy and will be presented below in Section 2 below.

The rest of the paper is devoted to the proof of sufficiency in Theorem 1.12.

2. THE CASE $1 < p \leq q$: THE SUFFICIENCY

Recall that as we explained above in Remark 1.11 a slightly more general statement that Theorem 1.10 was proved in [1]. The proof in this section is presented only for the reader's convenience.

Necessity of condition (1.13) in Theorem 1.10 was explained above in Section 1.3.3. Below we show that this condition is sufficient, i.e. that it implies estimate (1.12).

2.1. Construction of the stopping moments. The following standard construction is used for the proof of Theorem 1.10, it will also be needed later in the paper.

Given a (finite) collection $\mathcal{F} \subset \mathcal{L}$ of the intervals, we construct the collection $\mathcal{G} \subset \mathcal{F}$ of the stopping moments as follows. Given a non-negative function $f \geq 0$. For $J \in \mathcal{F}$, let $\mathcal{G}^*(J)$ be the collection of maximal intervals $I \in \mathcal{F}$, $I \subseteq J$ such that

$$\langle f \rangle_{I, \mu} > 2\langle f \rangle_{J, \mu}.$$

Note that intervals from $\mathcal{G}^*(J)$ are pairwise disjoint. Let $\mathcal{F}(J) = \{I \in \mathcal{F} : I \subseteq J\}$ and let $G(J) = \bigcup_{I \in \mathcal{G}^*(J)} I$. Define also $\mathcal{E}(J) = \mathcal{F}(J) \setminus \bigcup_{I \in \mathcal{G}^*(J)} \mathcal{F}(I)$. Then we have the following properties

- (i) For any $I \in \mathcal{E}(J)$, $\langle f \rangle_{I, \mu} \leq 2\langle f \rangle_{J, \mu}$,
- (ii) $\mu(G(J)) < \frac{1}{2}\mu(J)$.

To construct a collection \mathcal{G} , consider all maximal (by inclusion) intervals $J \in \mathcal{F}$. These sets form the first generation \mathcal{G}_1^* of stopping moments. Inductively define the $(n+1)$ -th generation

of stopping moments by $\mathcal{G}_{n+1}^* = \bigcup_{I \in \mathcal{G}_n^*} \mathcal{G}^*(I)$ and we define the collection of stopping moments by $\mathcal{G} = \bigcup_{n \geq 1} \mathcal{G}_n^*$.

Property (ii) implies that the collection \mathcal{G} of stopping moments satisfies the famous *Carleson measure condition* with constant 2

$$(2.1) \quad \sum_{I \in \mathcal{G}, I \subseteq J} \mu(I) < 2\mu(J), \quad J \in \mathcal{L}.$$

We will use the following well-known result.

Theorem 2.1 (Martingale Carleson Embedding Theorem). *Let μ be a measure on $(\mathcal{X}, \mathcal{T})$ and let $\alpha_I \geq 0$, $I \in \mathcal{L}$ satisfy the Carleson measure condition*

$$\sum_{I \in \mathcal{G}, I \subseteq J} \alpha_I \leq C\mu(J).$$

Then for any measurable function f and any $1 < p < \infty$

$$\sum_{I \in \mathcal{L}} \alpha_I \left| \langle f \rangle_{I, \mu} \right|^p \leq C \cdot (p')^p \|f\|_{L^p(\mathcal{X}, \mathcal{T}, \mu)}^p,$$

where $1/p' + 1/p = 1$.

This theorem with some constant instead of $(p')^p$ is a well known fact in harmonic analysis. The constant $(p')^p$ can be obtained, for example, from the fact that the norm of the martingale maximal function in L^p is estimated by p' . This connection with the constant in the Carleson Embedding Theorem is explained in Section 4 of [18]; as for the constant in estimate of the maximal function see, for example, [22, Theorem 14.1].

For a direct proof of Theorem 2.1 see [10]; optimality of the constant $(p')^p$ is also proved there.

2.2. Sufficiency in Theorem 1.8. First of all notice that it is sufficient to prove (1.9) only for $f \geq 0$, so let us assume without loss of generality that $f \geq 0$.

Using standard approximation reasoning we can also assume without loss of generality that only finitely many of the terms $\beta_{I, I'}$ are non-zero.

Then, applying the construction from the previous section with

$$\mathcal{F} = \{I \in \mathcal{L} : \exists I' \in \text{ch}(I) \text{ s.t. } \beta_{I, I'} \neq 0\}$$

we get the collection \mathcal{G} of the stopping moments.

For $J \in \mathcal{G}$ denote

$$F_J = \left(\sum_{I \in \mathcal{E}(J)} \left| \langle f \rangle_{I, \mu} \right|^q |b_I|^q \right)^{1/q},$$

so

$$\left(\sum_{I \in \mathcal{L}} \left| \langle f \rangle_{I, \mu} \right|^q |b_I|^q \right)^{p/q} = \left(\sum_{J \in \mathcal{G}} \sum_{I \in \mathcal{E}(J)} \left| \langle f \rangle_{I, \mu} \right|^q |b_I|^q \right)^{p/q} = \left(\sum_{J \in \mathcal{G}} |F_J|^q \right)^{p/q}$$

Note that

$$\begin{aligned}
 (2.2) \quad \int |F_J|^p d\nu &= \int_{\mathcal{X}} \left(\sum_{I \in \mathcal{E}(J)} |\langle f \rangle_{I,\mu}|^q |b_I|^q \right)^{p/q} d\nu \\
 &\leq 2^p |\langle f \rangle_{J,\mu}|^p \int_{\mathcal{X}} \left(\sum_{I \in \mathcal{E}(J)} |b_I|^q \right)^{p/q} d\nu && \text{by (i)} \\
 &\leq 2^p |\langle f \rangle_{J,\mu}|^p B^p \mu(J) && \text{by (1.13) .}
 \end{aligned}$$

Using the fact that $\|x\|_{\ell^q} \leq \|x\|_{\ell^p}$ for $p \leq q$, we can estimate

$$\begin{aligned}
 \int_{\mathcal{X}} \left(\sum_{I \in \mathcal{L}} |\langle f \rangle_{I,\mu}|^q |b_I|^q \right)^{p/q} d\nu &= \int_{\mathcal{X}} \left(\sum_{J \in \mathcal{G}} |F_J|^q \right)^{p/q} d\nu \\
 &\leq \int_{\mathcal{X}} \left(\sum_{J \in \mathcal{G}} |F_J|^p \right) d\nu && \text{because } \|x\|_{\ell^q} \leq \|x\|_{\ell^p} \\
 &\leq 2^p B^p \sum_{J \in \mathcal{G}} |\langle f \rangle_{J,\mu}|^p \mu(J) && \text{by (2.2)} \\
 &\leq 2^{p+1} (p')^p B^p \|f\|_{L^p(\mu)}^p && \text{by Theorem 2.1 and (2.1).}
 \end{aligned}$$

□

3. THE CASE $2 < p < \infty$: A COUNTEREXAMPLE

In this section we show that for $2 < p < \infty$ condition (1.5) alone is not sufficient for (1.4). More precisely, we construct a counterexample, showing that Theorem 1.8 with the special choice of β given in Remark 1.7 fails for any $p > 2$, i.e. that condition (1.10) does not imply (1.9) in this case. By the Littlewood–Paley estimates (Theorem 1.6) this will imply that condition (1.5) is not sufficient for (1.4) for any $p > 2$.

Consider the real line \mathbb{R} with the Borel σ -algebra $\mathcal{B}(\mathbb{R})$. Let the σ -algebra \mathfrak{S}_n be generated by the triadic intervals $I_k^n = 3^{-n}[k, k+1)$, $k \in \mathbb{Z}$, so \mathcal{L} is the collection of all triadic intervals I_k^n , $k, n \in \mathbb{Z}$.

We specify the measures μ, ν and the functions b, f in the following way.

Let C_n , $n \geq 0$ be the sets similar¹ to the ones appearing in the construction of the $\frac{1}{3}$ -Cantor set, namely $C_0 = [0, 1)$, $C_1 = [0, 1/3) \cup [2/3, 1)$ and, in general,

$$C_n = \bigcup_{x = \sum_{j=1}^n \varepsilon_j 3^{-j}, \varepsilon_j \in \{0, 2\}} [x, x + 3^{-n}).$$

Note that the above intervals $[x, x + 3^{-n})$ are the connected components of C_n , and C_n is a union of 2^n such intervals.

¹The $\frac{1}{3}$ -Cantor set is defined as $\bigcap_{n \geq 0} \overline{C}_n$, where \overline{C}_n is the closure of C_n .

The measure μ is taken to be the Lebesgue measure restricted on $[0, 1)$ and the measure ν is the Cantor measure, i.e. for each triadic interval $I = I_k^n = 3^{-n}[k, k+1)$ of length 3^{-n} we have $\nu(I) = 2^{-n}$ if $I \subset C_n$ and $\nu(I) = 0$ otherwise.

For the function b , we specify its martingale differences $\Delta_I^\nu b$. Let $I = I_k^n$ be a triadic interval, $I \subset C_n$, and let I_\pm be its right and left thirds respectively (note that $I_\pm \subset C_{n+1}$).

Assume that $p > 2$ is fixed, and define $\Delta_I^\nu := (2/3)^{n/p} (\mathbf{1}_{I_+} - \mathbf{1}_{I_-})$.

Now let us construct the function f . The set $D_n := C_{n-1} \setminus C_n$, $n \geq 1$ is a union of 2^{n-1} triadic intervals of length 3^{-n} , and the sets D_n are clearly disjoint. Fix r , $1/p < r < 1/2$ and define

$$f := \sum_{n \in \mathbb{N}} (3/2)^{n/p} n^{-r} \mathbf{1}_{D_n}.$$

Claim 3.1. The above construction gives a counterexample.

Proof. We first show that condition (1.10) with $b_I = \Delta_I^\nu$ is satisfied.

Let $J = I_k^n$ be a triadic interval of length 3^{-n} . If $J \not\subset C_n$, then $\pi_b^{\mu, \nu} \mathbf{1}_J = 0$ (equivalently, $\Pi_\beta \mathbf{1}_J = 0$), so we only need to check the case $J \subset C_n$.

Note that by the definition of the measures

$$\mu(J) = 3^{-n}, \quad \nu(J) = 2^{-n}.$$

We then can estimate

$$\left(\sum_{I \in \mathcal{L}(J)} |\Delta_I^\nu b|^2 \right)^{p/2} \leq \left(\sum_{k \geq n} (2/3)^{2k/p} \right)^{p/2} \leq B^p \left(\frac{2}{3} \right)^n,$$

where $B = (1 - (2/3)^{2/p})^{-1/2}$. Hence,

$$\int_J \left(\sum_{I \in \mathcal{L}(J)} |\Delta_I^\nu b|^2 \right)^{p/2} d\nu \leq B^p \left(\frac{2}{3} \right)^n \nu(J) = B^p \left(\frac{2}{3} \right)^n \cdot \left(\frac{1}{2} \right)^n = B^p \mu(J),$$

so (1.10) holds.

Next, we show that condition (1.9) fails. Recall that the sets D_n are disjoint, $\mu(D_n) = 2^{n-1}3^{-n}$ and that $f = (3/2)^{n/p} n^{-r}$ on D_n . So we can estimate

$$\|f\|_{L^p(\mu)}^p = \int_0^1 |f|^p dx = \sum_{n \geq 1} (3/2)^n n^{-pr} 2^{n-1} 3^{-n} = 2^{-1} \sum_{n \geq 1} n^{-pr} < \infty,$$

because $r > 1/p$ and so $rp > 1$.

Let us estimate $\sum_{I \in \mathcal{L}} |\langle f \rangle_{I, \mu}|^2 \cdot |\Delta_I^\nu b|^2$ on C_n . Take a triadic interval I of length 3^{-k} , $I \subset C_k$. Then the middle third of this interval is a subset of D_{k+1} , so

$$|\langle f \rangle_{I, \mu}| \geq 3^{-1} (3/2)^{(k+1)/p} (k+1)^{-r}.$$

We can see from the definition of $\Delta_I^\nu b$ that on $C_{k+1} \cap I$ (and so on $C_n \cap I$ for any $n > k$)

$$|\Delta_I^\nu b| = (2/3)^{k/p}.$$

so on $C_n \cap I$ we have

$$|\langle f \rangle_{I,\mu}| \cdot |\Delta_I^\nu b| \geq 3^{-1}(3/2)^{1/p}(k+1)^{-r}.$$

For a connected component I_n of C_n there exist connected components I_k of C_k , $k = 0, 1, \dots, n-1$ such that $I_n \subset I_k$. Therefore, squaring the above inequality and taking the sum over $k = 0, 1, \dots, n-1$ we get that on C_n .

$$\sum_{I \in \mathcal{L}} |\langle f \rangle_{I,\mu}|^2 |\Delta_I^\nu b|^2 \geq 3^{-2}(3/2)^{2/p} \sum_{k=1}^n k^{-2r}$$

Since $\nu(C_n) = 1$ we get that

$$\int_{C_n} \left(\sum_{I \in \mathcal{L}} |\langle f \rangle_{I,\mu}|^2 |\Delta_I^\nu b|^2 \right)^{p/2} d\nu \geq 3^{-p}(3/2)^{p/2} \left(\sum_{k=1}^n k^{-r} \right)^{p/2}.$$

The right hand side of this inequality tends to ∞ as $n \rightarrow \infty$, so (1.9) fails. \square

4. THE CASE $2 < p < \infty$: REDUCTION TO THE SHIFTED POSITIVE MARTINGALE OPERATORS AND NECESSITY OF CONDITION (1.14)

In this section we assume that q , $1 < q < \infty$ is fixed.

4.1. An auxiliary positive martingale operator. Consider an auxiliary non-linear operator $\Phi = \Phi_\beta = \Phi_\beta^q$,

$$\Phi_\beta f = \sum_{I \in \mathcal{L}} |\langle f \rangle_{I,\mu}|^q |b_I|^q;$$

since q is assumed to be fixed, we will skip index q to simplify the notation.

Estimate (1.12) is equivalent to the estimate

$$(4.1) \quad \|\Phi_\beta f\|_{L^{p/q}(\nu)}^{1/q} \leq A \|f\|_{L^p(\mu)} \quad \forall f \in L^p(\mu).$$

Consider also an auxiliary linear operator $\tilde{\Pi} = \tilde{\Pi}_\beta$,

$$\tilde{\Pi}_\beta f = \sum_{I \in \mathcal{L}} \langle f \rangle_{I,\mu} |b_I|^q$$

By Hölder inequality $|\langle f \rangle_{I,\mu}|^p \leq \langle |f|^p \rangle_{I,\mu}$, so $\Phi_\beta f \leq \tilde{\Pi}_\beta(|f|^q)$ (pointwise estimate). Therefore (1.12) follows from the estimate

$$(4.2) \quad \|\tilde{\Pi}_\beta g\|_{L^{p/q}(\nu)} \leq A^q \|g\|_{L^{p/q}(\mu)} \quad \forall g \in L^{p/q}(\mu)$$

(put $g = |f|^q$ in (4.1) and notice that $\|g\|_{L^{p/q}(\mu)} = \|f\|_{L^p(\mu)}^q$).

The operator $\tilde{\Pi}$ looks almost like the so-called *positive dyadic operators*: if the functions b_I were constants on I (recall that b_I are supported on I) we would get exactly a positive dyadic operator. But we only have that b_I is constant on children of I , so we say that the operator $\tilde{\Pi}_\beta$ is a *shifted* positive dyadic operator.

For positive dyadic operators the Sawyer type testing conditions are necessary and sufficient for the two weight estimates. If the same holds for the shifted positive martingale operators

(it is reasonable to expect, and we will prove it later in Section 5), we will get necessary and sufficient conditions for two weight estimate for paraproducts.

Let us write down the testing conditions for $\tilde{\Pi}_\beta$. Applying operator $\tilde{\Pi}_\beta$ to $\mathbf{1}_J$ and taking only terms corresponding to $I \subset J$ in the sum, we get the following necessary condition for the boundedness of $\tilde{\Pi}_\beta$

$$(4.3) \quad \int_J \left(\sum_{I \in \mathcal{L}(J)} |b_I|^q \right)^{p/q} d\nu \leq A^p \|\mathbf{1}_J\|_{L^{p/q}(\mu)}^{p/q} = A^p \mu(J)$$

Note that this is exactly the testing condition (1.13) with A instead of B . As we discussed above in Section 1.3.3 this condition is also necessary for the boundedness of the original operator Π_β .

Applying the adjoint of $\tilde{\Pi}_\beta$ to $\mathbf{1}_J$ and again taking in the sum only terms corresponding to $I \subset J$ and denoting by r' the Hölder conjugate to $r = p/q$, $1/r + 1/r' = 1$, we get another condition, necessary for the boundedness of $\tilde{\Pi}_\beta$,

$$(4.4) \quad \int_J \left(\sum_{I \in \mathcal{L}(J)} |b_I|^q \nu(I)/\mu(I) \right)^{r'} d\mu \leq A^{r'} \|\mathbf{1}_J\|_{L^{r'}(\nu)}^{r'} = A^{r'} \nu(J).$$

Note that this condition is exactly condition (1.14) with A instead of B_* . This condition is clearly necessary for the boundedness of Π_β . The proposition below implies that it is necessary for the boundedness of $\tilde{\Pi}_\beta$.

Proposition 4.1. *The operators $\Pi_\beta : L^p(\mu) \rightarrow \dot{\mathbf{g}}_p^q$ and $\tilde{\Pi}_\beta : L^r(\mu) \rightarrow L^r(\nu)$, $r = p/q$ are bounded simultaneously. Moreover,*

$$(4.5) \quad \|\Pi_\beta\|_{L^p(\mu) \rightarrow \dot{\mathbf{g}}_p^q(\nu)}^q \leq \|\tilde{\Pi}_\beta\|_{L^r(\mu) \rightarrow L^r(\nu)} \leq 4p(p-q)^{-1} \|\Pi_\beta\|_{L^p(\mu) \rightarrow \dot{\mathbf{g}}_p^q(\nu)}^q.$$

Let us first notice that the fact that the norms in (4.5) have different powers is not a typo: to see that the powers should be like that one just need to see what happens when we multiply the sequence β by a constant $a > 0$.

Note, that the first inequality in (4.5) was already proved, it is just the fact that (4.2) implies (4.1), so we only need to prove the second estimate.

4.2. Rubio de Francia operator and proof of Proposition 4.1. We need the well-known facts about the so-called Rubio de Francia operator. Recall that the martingale maximal function $M_\mu = M_{\mathcal{L}, \mu}$ is defined as

$$M_\mu f(x) = \sup_{I \in \mathcal{L}: x \in I} |\langle f \rangle_{I, \mu}|,$$

and that for $1 < p < \infty$

$$\|M_\mu f\|_{L^p(\mu)} \leq p' \|f\|_{L^p(\mu)};$$

here recall p' is the conjugate Hölder exponent, $1/p + 1/p' = 1$.

Let M_μ^n denotes the iterated maximal function, $M_\mu^2 f = M_\mu(M_\mu f)$, $M_\mu^{n+1} f = M_\mu(M_\mu^n f)$. The Rubio de Francia operator $R_p = R_{\mu, p}$, $1 < p < \infty$ is defined as

$$R_{\mu, p} f = \sum_{n \geq 0} (2p')^{-n} M_\mu^n f;$$

here $M_\mu^0 f = |f|$.

Proposition 4.2. *Let $f \geq 0$, $f \in L^p(\mu)$ Then*

- (i) $f \leq R_{\mu,p} f$;
- (ii) $\|R_{\mu,p} f\|_{L^p(\mu)} \leq 2\|f\|_{L^p(\mu)}$;
- (iii) For any $I \in \mathcal{L}$

$$\inf_{x \in I} R_{\mu,p} f(x) \geq (2p')^{-1} \langle R_{\mu,p} f \rangle_{I,\mu}.$$

The proof is well-known, see for example [2, Chapter IV, Lemma 5.1], and pretty straightforward. We leave it as an exercise for the reader.

Proof of Proposition 4.1. Denote $r = p/q$. For $f \in L^r(\mu)$ define

$$\tilde{f} := R_{\mu,r} |f|, \quad g := \tilde{f}^{1/q}.$$

By statement (ii) of Proposition 4.2

$$(4.6) \quad \|g\|_{L^p(\mu)}^q = \|\tilde{f}\|_{L^r(\mu)} \leq 2\|f\|_{L^r(\mu)}.$$

By statement (iii) of Proposition 4.2 for any $I \in \mathcal{L}$

$$(4.7) \quad \langle g^q \rangle_{I,\mu} = \langle \tilde{f} \rangle_{I,\mu} \leq 2r' \inf_{x \in I} \tilde{f}(x) = 2r' \inf_{x \in I} g(x)^q \leq 2r' \langle g \rangle_{I,\mu}^q;$$

the last inequality is the trivial fact that infimum is bounded by the average.

By statement (i) of Proposition 4.2 we have the a.e. estimate

$$(4.8) \quad |\tilde{\Pi}_\beta f| \leq \tilde{\Pi}_\beta |f| \leq \tilde{\Pi}_\beta \tilde{f}.$$

By the reverse Hölder inequality (4.7)

$$(4.9) \quad \tilde{\Pi}_\beta \tilde{f} \leq 2r' \sum_{I \in \mathcal{L}} |\langle g \rangle_{I,\mu}|^q |b_I|^q,$$

so, taking into account (4.8) we get

$$\begin{aligned} \|\tilde{\Pi}_\beta f\|_{L^r(\nu)} &\leq \|\tilde{\Pi}_\beta \tilde{f}\|_{L^r(\nu)} && \text{by (4.8)} \\ &\leq 2r' \|\Pi_\beta g\|_{\dot{\mathbf{g}}_p^q(\nu)}^q && \text{by (4.9)} \\ &\leq 2r' \|\Pi_\beta\|_{L^p(\mu) \rightarrow \dot{\mathbf{g}}_p^q(\nu)}^q \|g\|_{\dot{\mathbf{g}}_p^q(\nu)}^q \\ &\leq 4r' \|\Pi_\beta\|_{L^p(\mu) \rightarrow \dot{\mathbf{g}}_p^q(\nu)}^q \|f\|_{L^r(\mu)} && \text{by (4.6)}, \end{aligned}$$

so the second inequality in (4.5) holds.

And as we discussed immediately after Proposition 4.1, the first inequality in (4.5) is already proved. \square

Remark. Note that for $p \rightarrow q^+$ the constants blow up. Namely, from Proposition 4.1 we get that the constant B_* in the dual testing condition (1.14) can be estimated as $B_* \leq c(p,q)^{-1} A^q = 4r' A^q$, where A is the estimate of the norm of Π_β . Note that $c(p,q) \rightarrow \infty$ as $p \rightarrow q^+$, which means that using our approach we cannot prove the necessity of the dual testing condition (1.14) in the limit case $p = q$. And in fact, the limiting case $p = q$ of the

condition (1.14) is not necessary: this can be seen, for example, in the simplest case when \mathcal{L} is the standard dyadic lattice \mathcal{D} on \mathbb{R} , $p = q = 2$ and $\mu = \nu$ is the Lebesgue measure.

4.3. Bilinear operators and trilinear forms. The discussion in this subsection is not necessary of the proof of the main result, but it might be of independent interest.

Estimate (4.1) for $q = 2$ can be interpreted as an estimate of the bilinear operator $L_\beta : L^p(\mu) \times L^p(\mu) \rightarrow L^{p/2}(\nu)$,

$$(4.10) \quad L_\beta(f, g) = \sum_{I \in \mathcal{L}} \langle f \rangle_{I, \mu} \langle g \rangle_{I, \mu} |b_I|^2.$$

Indeed, by the definition (4.10) of L_β

$$|L_\beta(f, g)| \leq \frac{1}{2} \left(\Phi_\beta(f) + \Phi_\beta(g) \right),$$

so if (4.1) is satisfied then

$$|L_\beta(f, g)| \leq A^2 \frac{1}{2} \left(\|f\|_{L^p(\mu)}^2 + \|g\|_{L^p(\mu)}^2 \right).$$

Replacing in this inequality f by tf and g by $t^{-1}g$ (note that this does not change the left hand side) and taking the infimum over all $t > 0$ in the right hand side we get that

$$(4.11) \quad |L_\beta(f, g)| \leq A^2 \|f\|_{L^p(\mu)} \|g\|_{L^p(\mu)}.$$

On the other hand, if (4.11) holds, then plugging in $|f|$ for f and g we get exactly (4.1).

4.4. Reduction to the estimates of shifted positive martingale operators. We reduced Theorem 1.12 to the estimates of the auxiliary shifted positive martingale operator $\tilde{\Pi}_\beta$.

Let us write this operator in more symmetric form. Recall that for $I \in \mathcal{L}$ the symbol \hat{I} denotes the parent of I .

For a sequence $\{\alpha_{\hat{I}, I}\}_{I \in \mathcal{L}}$, $\alpha_{\hat{I}, I} \geq 0$ define the operator T_α ,

$$(4.12) \quad T_\alpha f = \sum_{I \in \mathcal{L}} \left(\int_I f d\mu \right) a_I, \quad \text{where } a_I = \sum_{I' \in \text{ch}(I)} \alpha_{\hat{I}, I'} \mathbf{1}_{I'}$$

Note that if $\alpha_{\hat{I}, I'} = \mu(I)^{-1} |\beta_{I, I'}|^p$, so $a_I = \mu(I)^{-1} |b_I|^p$, then T_α is exactly the operator $\tilde{\Pi}_\beta$.

Theorem 1.9 follows from the theorem below, applied with $r = p/q$ instead of p .

Theorem 4.3 (Two weight estimates for shifted positive operator). *Let $\alpha = \{\alpha_{\hat{I}, I}\}_{I \in \mathcal{L}}$ be a sequence of non-negative constants and let T_α be the operator defined by (4.12)*

Then

$$\|T_\alpha f\|_{L^p(\nu)} \leq A \|f\|_{L^p(\mu)} \quad \forall f \in L^p(\mu)$$

if and only if

$$(4.13) \quad \int_J \left(\sum_{I \in \mathcal{L}(J)} \mu(I) a_I \right)^p d\nu \leq B^p \mu(J),$$

and

$$(4.14) \quad \int_J \left(\sum_{I \in \mathcal{L}(J)} \nu(I) a_I \right)^{p'} d\mu \leq B_*^{p'} \nu(J).$$

Moreover, for the best constants we have $\max\{B, B_*\} \leq A \leq C(p) \max\{B, B_*\}$

Remark. We can see from Proposition 4.5 that constants in the above Theorem 4.3 (with $r = p/q$ instead of p) are q th powers of the constants in Theorem 1.12. Note also that constant A in (4.3) is the same as in Theorem 1.12.

5. TWO WEIGHT ESTIMATES FOR SHIFTED POSITIVE MARTINGALE OPERATORS

In this section we prove Theorem 4.3, and hence Theorem 1.12.

Necessity of the conditions (4.13) and (4.14) and the estimates $B, B_* \leq A$ are trivial: we just test operators T_α and its formal adjoint on functions $\mathbf{1}_J$ and count only part of the sum. This necessity was already discussed in the previous section, see (4.3), (4.4).

So we only need to prove the sufficiency. Rescaling α we can assume without loss of generality that $B = B_* = 1$ (the constants are not assumed to be optimal).

The idea of the proof comes from [18], but requires some new ideas. Namely, positive martingale operators treated in [18] are the special case of our operator T_α where all a_I are constant on I , $a_I = \alpha_I \mathbf{1}_I$, and this property was essential in the proof in [18].

In this paper we only know that a_I are constant on children of I , so the proof requires modification of the standard stopping moment construction.

Take $f \geq 0$ and $g \geq 0$, $\|f\|_{L^p(\mu)} = \|g\|_{L^{p'}(\nu)} = 1$. To prove sufficiency in Theorem 4.3 it is enough to show that

$$(5.1) \quad \langle T_\alpha f, g \rangle_\nu = \sum_{I \in \mathcal{L}} \alpha_{\widehat{I}, I} \left(\int_{\widehat{I}} f d\mu \right) \left(\int_I g d\nu \right) \leq A.$$

We split the above sum into two parts by splitting $\mathcal{L} = \mathcal{A} \cup \mathcal{B}$ according to the following splitting condition:

$$(5.2) \quad \mathcal{A} = \left\{ I \in \mathcal{L} : \langle f \rangle_{I, \mu}^p \cdot \mu(I) \geq \langle g \rangle_{I, \nu}^{p'} \cdot \nu(I) \right\} \text{ and } \mathcal{B} = \mathcal{L} \setminus \mathcal{A}.$$

Standard approximation reasoning allows us to assume without loss of generality that only finitely many terms $\alpha_{\widehat{I}, I}$ are non-zero, so all the sums are finite.

Using the splitting condition (5.2), we can write $\langle T_\alpha f, g \rangle_\nu = T_1 + T_2$, where

$$(5.3) \quad T_1 = \sum_{I \in \mathcal{A}} \alpha_{\widehat{I}, I} \left(\int_{\widehat{I}} f d\mu \right) \left(\int_I g d\nu \right),$$

$$(5.4) \quad T_2 = \sum_{I \in \mathcal{B}} \alpha_{\widehat{I}, I} \left(\int_{\widehat{I}} f d\mu \right) \left(\int_I g d\nu \right).$$

5.1. A modified stopping interval construction. To estimate T_1 we need to modify a bit the construction of stopping intervals from Section 2.1. The main feature of the construction is that the stopping intervals will be the intervals $I \in \mathcal{A}$, but the stopping criterion will be checked on their parents \widehat{I} .

We start with some interval J (not necessarily in \mathcal{A}). For the interval J we define the primary *preliminary stopping intervals* to be the maximal by inclusion intervals $\widehat{I} \subset J$, $I \in \mathcal{A}$, such that

$$(5.5) \quad \langle f \rangle_{\widehat{I}, \mu} \geq 2 \langle f \rangle_{J, \mu};$$

note that different $I \in \mathcal{A}$ can give the same \widehat{I} , but this \widehat{I} is counted only once.

It is obvious that these preliminary stopping intervals are disjoint and their total μ -measure is at most $\mu(J)/2$.

For each such preliminary stopping interval pick all its children L that belong to \mathcal{A} (there is at least one such L), and declare these children to be the stopping intervals.

For the children $K \notin \mathcal{A}$ we continue the process: we will find the maximal by inclusion intervals $\widehat{I} \subset K$, $I \in \mathcal{A}$ satisfying (5.5), and declare these \widehat{I} to be the secondary preliminary stopping intervals (note that in the stopping criterion (5.5) we still compare with the average over the original interval J).

For these preliminary stopping intervals we add their children $L \in \mathcal{A}$ to the stopping intervals, and for the children $K \notin \mathcal{A}$ we continue the process (still comparing the averages with the average over the original interval J).

We assumed that the collection \mathcal{A} is finite, so at some point the process will stop (no $I \in \mathcal{A}$, $\widehat{I} \subset K$). We end up with the disjoint collection $\mathcal{G}^*(J) = \mathcal{G}_{\mathcal{A}}^*(J)$ of stopping intervals.

Since all the stopping intervals are inside the *primary* preliminary stopping intervals, we can conclude that

$$(5.6) \quad \sum_{I \in \mathcal{G}^*(J)} \mu(I) \leq \frac{1}{2} \mu(J).$$

Denote $G(J) := \bigcup_{I \in \mathcal{G}^*(J)} I$. Denoting

$$\mathcal{A}(J) := \{I \in \mathcal{A} : I \subset J\}, \quad \mathcal{A}'(J) := \mathcal{A}(J) \setminus \{J\},$$

define

$$\mathcal{E}(J) = \mathcal{E}_{\mathcal{A}}(J) := \mathcal{A}(J) \setminus \bigcup_{I \in \mathcal{G}^*(J)} \mathcal{A}(I), \quad \mathcal{E}'(J) := \mathcal{E}(J) \setminus \{J\}.$$

It easily follows from the construction that for any $I \in \mathcal{E}'(J)$

$$(5.7) \quad \langle f \rangle_{\widehat{I}, \mu} \leq 2 \langle f \rangle_{J, \mu}.$$

5.2. Estimate of T_1 . To estimate T_1 we run the stopping moments construction defined above in Section 5.1. We start with the collection \mathcal{G}_0 of disjoint intervals covering the set $\bigcup_{I \in \mathcal{A}} \widehat{I}$. Note that then $\mathcal{G}_0 \cap \mathcal{A} = \emptyset$.

For each $I \in \mathcal{G}_0$ we run the stopping moments construction to get the collection $\mathcal{G}^*(I)$; the union $\bigcup_{I \in \mathcal{G}_0} \mathcal{G}^*(I)$ give us the first generation of stopping moments \mathcal{G}_1^* . Define inductively

$$\mathcal{G}_{k+1}^* := \bigcup_{I \in \mathcal{G}_k^*} \mathcal{G}^*(I),$$

and put $\mathcal{G} := \bigcup_{k \geq 1} \mathcal{G}_k^*$.

Note that the condition (5.6) implies that the collection \mathcal{G} satisfies the following Carleson measure condition

$$(5.8) \quad |I_0|^{-1} \sum_{I \in \mathcal{G}: I \subset I_0} \mu(I) \leq 2 \quad \forall I_0 \in \mathcal{L};$$

we also can replace \mathcal{G} by $\mathcal{G} \cup \mathcal{G}_0$ here, and still have the same estimate.

Since the collection \mathcal{A} is the disjoint union of the collections $\mathcal{E}'(I)$, $I \in \mathcal{G} \cup \mathcal{G}_0$ and the collection \mathcal{G} , we can represent T_1 as $T_1 = S_1 + S_2$, where

$$S_1 = \sum_{J \in \mathcal{G} \cup \mathcal{G}_0} \int_J F_J g \, d\nu,$$

with

$$(5.9) \quad F_J := \sum_{I \in \mathcal{E}'(J)} \left(\int_{\hat{I}} f \, d\mu \right) a_{\hat{I}, I} \mathbf{1}_I,$$

and

$$(5.10) \quad S_2 = \sum_{I \in \mathcal{G}} a_{\hat{I}, I} \left(\int_{\hat{I}} f \, d\mu \right) \left(\int_I g \, d\nu \right).$$

Note, that as defined, F_J is supported by J .

Let us estimate S_1 . The estimate (5.7) of the average together with the testing condition (4.13) imply that

$$(5.11) \quad \|F_J\|_{L^p(\nu)} \leq 2 \langle f \rangle_{J, \mu} \mu(J)^{1/p}$$

Let us write

$$\begin{aligned} S_1 &= \sum_{J \in \mathcal{G} \cup \mathcal{G}_0} \int_J F_J g \, d\nu \\ &= \sum_{J \in \mathcal{G} \cup \mathcal{G}_0} \int_{J \setminus G(J)} F_J g \, d\nu + \sum_{J \in \mathcal{G} \cup \mathcal{G}_0} \int_{G(J)} F_J g \, d\nu \\ &=: \sum_{J \in \mathcal{G} \cup \mathcal{G}_0} A(J) + \sum_{J \in \mathcal{G} \cup \mathcal{G}_0} B(J). \end{aligned}$$

The first sum is easy to estimate. By Hölder inequality and (5.11)

$$|A(J)| \leq \|F_J\|_{L^p(\nu)} \left(\int_{J \setminus G(J)} |g|^{p'} \, d\nu \right)^{1/p'}$$

$$\leq 2\langle f \rangle_{J,\mu} \mu(J)^{1/p} \left(\int_{J \setminus G(J)} |g|^{p'} d\nu \right)^{1/p'}$$

Then using again Hölder inequality and then the fact that the sets $J \setminus G(J)$, $J \in \mathcal{G} \cup \mathcal{G}_0$ are disjoint we get

$$\begin{aligned} \sum_{J \in \mathcal{G} \cup \mathcal{G}_0} A(J) &\leq 2 \left(\sum_{J \in \mathcal{G} \cup \mathcal{G}_0} \langle f \rangle_{J,\mu}^p \mu(J) \right)^{1/p} \left(\sum_{J \in \mathcal{G} \cup \mathcal{G}_0} \int_{J \setminus G(J)} |g|^{p'} d\nu \right)^{1/p'} \quad \text{by Hölder,} \\ &\leq 2 \left(\sum_{J \in \mathcal{G} \cup \mathcal{G}_0} \langle f \rangle_{J,\mu}^p \mu(J) \right)^{1/p} \|g\|_{L^{p'}(\nu)} \quad \text{sets } J \setminus G(J) \text{ are disjoint.} \end{aligned}$$

Since the collection $\mathcal{G} \cup \mathcal{G}_0$ satisfies the Carleson measure condition (5.8), we can apply the Carleson Embedding Theorem (Theorem 2.1) to get

$$(5.12) \quad \sum_{J \in \mathcal{G} \cup \mathcal{G}_0} \langle f \rangle_{J,\mu}^p \mu(J) \leq 2(p')^p \|f\|_{L^p(\mu)}^p,$$

so

$$\sum_{J \in \mathcal{G} \cup \mathcal{G}_0} A(J) \leq 2^{1+1/p} p' \|f\|_{L^p(\mu)} \|g\|_{L^{p'}(\nu)}.$$

To estimate $\sum B(J)$ we notice, that since the functions F_J are constant on each interval $I \in \mathcal{G}^*(J)$,

$$\int_{G(J)} F_J g d\nu = \int_{G(J)} F_J g_J d\nu,$$

where

$$g_J := \sum_{I \in \mathcal{G}^*(J)} \langle g \rangle_{I,\nu} \mathbf{1}_I$$

Then by Hölder inequality

$$\begin{aligned} B(J) &\leq \|F_J\|_{L^p(\nu)} \|g_J\|_{L^{p'}(\nu)} \\ &\leq 2\langle f \rangle_{J,\mu} \mu(J)^{1/p} \left(\sum_{I \in \mathcal{G}^*(J)} \langle g \rangle_{I,\nu}^{p'} \nu(I) \right)^{1/p'}. \end{aligned}$$

The stopping intervals $I \in \mathcal{G}^*(J)$ belong to \mathcal{A} , so by (5.2)

$$\langle g \rangle_{I,\nu}^{p'} \nu(I) \leq \langle f \rangle_{I,\mu}^p \mu(I),$$

and therefore

$$B(J) \leq 2\langle f \rangle_{J,\mu} \mu(J)^{1/p} \left(\sum_{I \in \mathcal{G}^*(J)} \langle f \rangle_{I,\mu}^p \mu(I) \right)^{1/p'}.$$

Summing and applying Hölder inequality we get

$$\begin{aligned} \sum_{J \in \mathcal{G} \cup \mathcal{G}_0} B(J) &\leq 2 \left(\sum_{J \in \mathcal{G} \cup \mathcal{G}_0} \langle f \rangle_{J, \mu}^p \mu(J) \right)^{1/p} \left(\sum_{J \in \mathcal{G}} \langle f \rangle_{I, \mu}^p \mu(I) \right)^{1/p'} \\ &\leq 2 \sum_{J \in \mathcal{G} \cup \mathcal{G}_0} \langle f \rangle_{J, \mu}^p \mu(J) \\ &\leq 4(p')^p \|f\|_{L^p(\mu)}^p \quad \text{by (5.12)} \end{aligned}$$

Recall that we assumed that $\|f\|_{L^p(\mu)} = \|g\|_{L^{p'}(\nu)} = 1$, so

$$\sum_{J \in \mathcal{G} \cup \mathcal{G}_0} B(J) \leq 4(p')^p \|f\|_{L^p(\mu)} \|g\|_{L^{p'}(\nu)},$$

and thus

$$S_1 \leq (4(p')^p + 2^{1+1/p} p') \|f\|_{L^p(\mu)} \|g\|_{L^{p'}(\nu)}.$$

To estimate S_2 , let us write recalling the splitting condition (5.2)

$$\begin{aligned} S_2 &= \sum_{I \in \mathcal{G}} a_{\widehat{I}, I} \langle f \rangle_{\widehat{I}, \mu} \mu(\widehat{I}) \langle g \rangle_{I, \nu} \nu(I) \\ &\leq \left(\sum_{I \in \mathcal{G}} \langle f \rangle_{\widehat{I}, \mu}^p a_{\widehat{I}, \mu}^p \mu(\widehat{I})^p \nu(I) \right)^{1/p} \left(\sum_{I \in \mathcal{G}} \langle g \rangle_{I, \nu}^{p'} \nu(I) \right)^{1/p'} \quad \text{Hölder inequality} \\ &\leq \left(\sum_{I \in \mathcal{G}} \langle f \rangle_{\widehat{I}, \mu}^p a_{\widehat{I}, \mu}^p \mu(\widehat{I})^p \nu(I) \right)^{1/p} \left(\sum_{I \in \mathcal{G}} \langle f \rangle_{I, \mu}^p \mu(I) \right)^{1/p'} \quad \text{by (5.2)} \end{aligned}$$

The second term is easy to estimate: the Carleson measure condition (5.8) and the Carleson Embedding Theorem (Theorem 2.1) imply that

$$\sum_{I \in \mathcal{G}} \langle f \rangle_{I, \mu}^p \mu(I) \leq 2(p')^p \|f\|_{L^p(\mu)}^p.$$

To estimate the first term, we apply Lemma 5.1 below.

Lemma 5.1. *The sequence $\{\alpha_I\}_{I \in \mathcal{L}}$,*

$$\alpha_I := \sum_{I' \in \text{ch}(I)} a_{I, I'}^p \mu(I)^p \nu(I')$$

satisfies the Carleson measure condition

$$(5.13) \quad \sum_{I \in \mathcal{L}: I \subset I_0} \alpha_I \leq \mu(I_0) \quad \forall I_0 \in \mathcal{L}.$$

Using the lemma we can estimate

$$\begin{aligned} (5.14) \quad \sum_{I \in \mathcal{G}} \langle f \rangle_{\widehat{I}, \mu}^p a_{\widehat{I}, \mu}^p \mu(\widehat{I})^p \nu(I) &\leq \sum_{I \in \mathcal{L}} \langle f \rangle_I^p \alpha_I \quad \text{sum over a bigger set} \\ &\leq (p')^p \|f\|_{L^p(\mu)}^p \quad \text{by Theorem 2.1.} \end{aligned}$$

Gathering the above estimates, we get the desired estimate of S_2 and therefore of T_1 . \square

Proof of Lemma 5.1. For $I_0 \in \mathcal{L}$

$$\begin{aligned} \sum_{I \in \mathcal{L}(I_0)} \alpha_I &= \int_{I_0} \left(\sum_{I \in \mathcal{L}(I_0)} \sum_{I' \in \text{ch}(I)} a_{I,I'}^p \mu(I)^p \mathbf{1}_{I'} \right) d\nu \\ &\leq \int_{I_0} \left(\sum_{I \in \mathcal{L}(I_0)} \sum_{I' \in \text{ch}(I)} a_{I,I'} \mu(I) \mathbf{1}_{I'} \right)^p d\nu && \text{because } \|x\|_{\ell^p} \leq \|x\|_{\ell^1} \\ &\leq \mu(I_0) && \text{by assumption (4.13).} \end{aligned}$$

\square

5.3. Estimate of T_2 . The estimate of T_2 is simpler, because it relies on a simpler stopping moment construction.

Namely, we run the stopping intervals construction described in Section 2.1 with $\mathcal{F} = \mathcal{B}$ and with respect to the measure ν and the function g .

We start with a collection \mathcal{G}_0 of disjoint intervals covering the set $\bigcup_{I \in \mathcal{B}} I$, and run the construction starting from these intervals. We get the collection \mathcal{G} of stopping intervals satisfying the Carleson measure condition

$$(5.15) \quad \sum_{I \in \mathcal{G} \cup \mathcal{G}_0, I \subset I_0} \nu(I) \leq 2\nu(I_0) \quad \forall I_0 \in \mathcal{L}.$$

Again, define $G(J) := \bigcup_{I \in \mathcal{G}^*(J)} I$. Denoting

$$\mathcal{B}(J) := \{I \in \mathcal{B} : I \subset \mathcal{B}\}, \quad \mathcal{B}'(J) := \{I \in \mathcal{B} : I \subsetneq \mathcal{B}\} = \mathcal{B}(J) \setminus \{J\},$$

define

$$\mathcal{E}(J) = \mathcal{E}_{\mathcal{B}}(J) := \mathcal{B}(J) \setminus \bigcup_{I \subset \mathcal{G}^*(J)} \mathcal{B}(J), \quad \mathcal{E}'(J) = \mathcal{E}'_{\mathcal{B}}(J) = \mathcal{E}(J) \setminus \{J\}.$$

Similarly to the case of T_1 we split the sum into 2 parts, $T_2 = S_1 + S_2$, where S_2 is the sum over \mathcal{G} and S_1 is the rest,

$$T_2 = \sum_{I \in \mathcal{B}: I \notin \mathcal{G}} a_{\hat{I},I} \left(\int_{\hat{I}} f d\mu \right) \left(\int_I g d\nu \right) + \sum_{I \in \mathcal{G}} a_{\hat{I},I} \left(\int_{\hat{I}} f d\mu \right) \left(\int_I g d\nu \right) =: S_1 + S_2.$$

Denoting for $J \in \mathcal{G} \cup \mathcal{G}_0$

$$F_J := \sum_{I \in \mathcal{E}'(J)} a_{\hat{I},I} \left(\int_I g d\nu \right) \mathbf{1}_{\hat{I}}$$

and noticing that F_J is supported on J we can write

$$S_1 = \sum_{J \in \mathcal{G} \cup \mathcal{G}_0} \int_J F_J f d\mu.$$

Using the estimate (5.15) and the testing condition (4.14) we can write

$$(5.16) \quad \|F_J\|_{L^{p'}(\mu)} \leq 2\langle g \rangle_{J,\nu} \nu(J)^{1/p'}$$

We then decompose S_1 as

$$\begin{aligned} S_1 &= \sum_{J \in \mathcal{G} \cup \mathcal{G}_0} \int_J F_J f \, d\mu \\ &= \sum_{J \in \mathcal{G} \cup \mathcal{G}_0} \int_{J \setminus G(J)} F_J f \, d\mu + \sum_{J \in \mathcal{G} \cup \mathcal{G}_0} \int_{G(J)} F_J f \, d\mu \\ &=: \sum_{J \in \mathcal{G} \cup \mathcal{G}_0} A(J) + \sum_{J \in \mathcal{G} \cup \mathcal{G}_0} B(J). \end{aligned}$$

The first sum is easy to estimate. By Hölder inequality and (5.16)

$$\begin{aligned} |A(J)| &\leq \|F_J\|_{L^{p'}(\mu)} \left(\int_{J \setminus G(J)} |f|^p d\mu \right)^{1/p} \\ &\leq 2\langle g \rangle_{J,\nu} \nu(J)^{1/p'} \left(\int_{J \setminus G(J)} |f|^p d\mu \right)^{1/p} \end{aligned}$$

Then using again Hölder inequality and then the fact that the sets $J \setminus G(J)$, $J \in \mathcal{G} \cup \mathcal{G}_0$ are disjoint we get

$$\begin{aligned} \sum_{J \in \mathcal{G} \cup \mathcal{G}_0} A(J) &\leq 2 \left(\sum_{J \in \mathcal{G} \cup \mathcal{G}_0} \langle g \rangle_{J,\nu}^{p'} \nu(J) \right)^{1/p'} \left(\sum_{J \in \mathcal{G} \cup \mathcal{G}_0} \int_{J \setminus G(J)} |f|^p d\mu \right)^{1/p} \quad \text{by Hölder,} \\ &\leq 2 \left(\sum_{J \in \mathcal{G} \cup \mathcal{G}_0} \langle g \rangle_{J,\nu}^{p'} \nu(J) \right)^{1/p'} \|f\|_{L^p(\mu)} \quad \text{sets } J \setminus G(J) \text{ are disjoint.} \end{aligned}$$

Since the collection $\mathcal{G} \cup \mathcal{G}_0$ satisfies the Carleson measure condition (5.15), we can apply the Carleson Embedding Theorem (Theorem 2.1) to get

$$(5.17) \quad \sum_{J \in \mathcal{G} \cup \mathcal{G}_0} \langle g \rangle_{J,\nu}^{p'} \nu(J) \leq 2p^{p'} \|g\|_{L^{p'}(\nu)}^{p'},$$

so

$$\sum_{J \in \mathcal{G} \cup \mathcal{G}_0} A(J) \leq 2^{1+1/p'} p \|f\|_{L^p(\mu)} \|g\|_{L^{p'}(\nu)}.$$

To estimate $\sum B(J)$ we notice, that since the functions F_J are constant on each interval $I \in \mathcal{G}^*(J)$,

$$\int_{G(J)} F_J f \, d\mu = \int_{G(J)} F_J f_J \, d\mu,$$

where

$$f_J := \sum_{I \in \mathcal{G}^*(J)} \langle f \rangle_{I,\mu} \mathbf{1}_I$$

Then by Hölder inequality

$$B(J) \leq \|F_J\|_{L^{p'}(\mu)} \|f_J\|_{L^p(\mu)}$$

$$\leq 2\langle g \rangle_{J,\nu} \nu(J)^{1/p'} \left(\sum_{I \in \mathcal{G}^*(J)} \langle f \rangle_{I,\mu}^p \mu(I) \right)^{1/p}.$$

The stopping intervals $I \in \mathcal{G}^*(J)$ belong to \mathcal{B} , so by (5.2)

$$\langle f \rangle_{I,\mu}^p \mu(I) < \langle g \rangle_{I,\nu}^{p'} \nu(I)$$

and therefore

$$B(J) \leq 2\langle g \rangle_{J,\nu} \nu(J)^{1/p'} \left(\sum_{I \in \mathcal{G}^*(J)} \langle g \rangle_{I,\nu}^{p'} \nu(I) \right)^{1/p}.$$

Summing and applying Hölder inequality we get

$$\begin{aligned} \sum_{J \in \mathcal{G} \cup \mathcal{G}_0} B(J) &\leq 2 \left(\sum_{J \in \mathcal{G} \cup \mathcal{G}_0} \langle g \rangle_{J,\nu}^{p'} \nu(J) \right)^{1/p'} \left(\sum_{J \in \mathcal{G}} \langle g \rangle_{J,\nu}^{p'} \nu(J) \right)^{1/p} \\ &\leq 2 \sum_{J \in \mathcal{G} \cup \mathcal{G}_0} \langle g \rangle_{J,\nu}^{p'} \nu(J) \\ &\leq 4p^{p'} \|g\|_{L^{p'}(\nu)}^{p'} \quad \text{by (5.17)} \end{aligned}$$

Recall that we assumed that $\|f\|_{L^p(\mu)} = \|g\|_{L^{p'}(\nu)} = 1$, so

$$\sum_{J \in \mathcal{G} \cup \mathcal{G}_0} B(J) \leq 4p^{p'} \|f\|_{L^p(\mu)} \|g\|_{L^{p'}(\nu)},$$

and thus

$$S_1 \leq (4p^{p'} + 2^{1+1/p'} p) \|f\|_{L^p(\mu)} \|g\|_{L^{p'}(\nu)}.$$

To estimate S_2 , let us write using Hölder inequality

$$\begin{aligned} S_2 &= \sum_{I \in \mathcal{G}} a_{\widehat{I},I} \langle f \rangle_{\widehat{I},\mu} \mu(\widehat{I}) \langle g \rangle_{I,\nu} \nu(I) \\ &\leq \left(\sum_{I \in \mathcal{G}} \langle f \rangle_{\widehat{I},\mu}^p a_{\widehat{I},\mu}^p \mu(\widehat{I})^p \nu(I) \right)^{1/p} \left(\sum_{I \in \mathcal{G}} \langle g \rangle_{I,\nu}^{p'} \nu(I) \right)^{1/p'} \end{aligned}$$

The second term was already estimated in (5.17). The first term is estimated using Lemma 5.1, exactly as it was done in the end of Section 5.2, see (5.14).

REFERENCES

- [1] A. Culiuc *A note on two weight bounds for the generalized Hardy-Littlewood Maximal operator*, arXiv:1506.07125 [math.CA] (215), 6 pp.
- [2] J. García-Cuerva, J. L. Rubio de Francia, *Weighted norm inequalities and related topics*, North-Holland Mathematics Studies, vol. 116, North-Holland Publishing Co., Amsterdam, 1985, Notas de Matemática [Mathematical Notes], 104.
- [3] T. Hänninen, *Another proof of Scurry's characterization of a two weight norm inequality for a sequence-valued positive dyadic operator*, ArXiv:1304.7759 [math.CA] (2013).

- [4] T. Hytönen, *The A_2 theorem: remarks and complements*, ArXiv:1212.3840 [math.CA] (2012).
- [5] M. Lacey, *Two Weight Inequality for the Hilbert Transform: A Real Variable Characterization II*, ArXiv:1301.4663 [math.CA] (2013), to appear in Duke Math. J.
- [6] M. Lacey, E. Sawyer and I. Uriarte-Tuero, *Two weight inequalities for discrete positive operators*, ArXiv:0911.3437 [math.CA] (2009).
- [7] M. Lacey, E. Sawyer, C.Y. Shen and I. Uriarte-Tuero, *Two Weight Inequality for the Hilbert Transform: A Real Variable Characterization I*, ArXiv:1201.4319 [math.CA] (2012), to appear in Duke Math. J.
- [8] M. Lacey, E. Sawyer, C.Y. Shen, I. Uriarte-Tuero and B. Wick, *Two Weight Inequalities for the Cauchy Transform from \mathbb{R} to \mathbb{C}_+* , ArXiv:1310.4820 [math.CV] (2013).
- [9] M. Lacey and B. Wick, *Two Weight Inequalities for Riesz Transforms: Uniformly Full Dimension Weights*, ArXiv:1312.6163 [math.CA] (2013).
- [10] J. Lai, *The Bellman functions of the Carleson Embedding Theorem and the Doob's martingale inequality*, arXiv:1411.5408 [math.CA], 2014.
- [11] F. Nazarov, S. Treil and A. Volberg, *The Bellman function and two weight inequalities for Haar multipliers*, J. Amer. Math. Soc. **12** (1999), no. 4, 909-928.
- [12] F. Nazarov, S. Treil and A. Volberg, *Two weight inequalities for individual Haar multipliers and other well localized operators*, Math. Res. Lett. **15** (2008), no. 3, 583-597.
- [13] J. Scurry, *A characterization of two-weight inequalities for a vector-valued operator*, ArXiv:1007.3089 [math.CA] (2010).
- [14] E. Sawyer, *A characterization of a two-weight norm inequality for maximal operators*, Studia Math. **75** (1982), no. 1, 1-11.
- [15] ———, *A characterization of two weight norm inequalities for fractional and Poisson integrals*, Trans. Amer. Math. Soc. **308** (1988), no. 2, 533-545.
- [16] E. Sawyer and R. Wheeden, *Weighted inequalities for fractional integrals on Euclidean and homogeneous spaces*, Amer. J. Math. **114** (1992), no. 4, 813-874.
- [17] H. Tanaka and Y. Terasawa *Positive operators and maximal operators in a filtered measure space*, J. Funct. Anal. **264** (2013), issue 4, 920-946.
- [18] S. Treil, *A remark on two weight estimates for positive dyadic operators*, in: K. Grchenig, Yu. Lyubarskii, K. Seip (Editors), *Operator-Related Function Theory and Time-Frequency Analysis: The Abel Symposium 2012 (Abel Symposia 9)*, Springer, 2014, p. 185-195; see also ArXiv:1201.1145 [math.CA] (2012).
- [19] S. Treil, *Commutators, paraproducts and BMO in non-homogeneous martingale settings*, Rev. Mat. Iberoam. **29** (2013), no. 4, 1325-1372, see also arXiv:1007.1210 [math.CA] (2010), 39 pp.
- [20] E. Vuorinen, *$L^p(\mu)$ - $L^q(\nu)$ characterization for well localized operators*, arXiv:1412.2127 [math.CA], 2014.
- [21] E. Vuorinen, *Two weight L^p -inequalities for dyadic shifts and the dyadic square function*, arXiv:1504.05759 [math.CA], 2015.
- [22] D. WILLIAMS, *Probability with martingales*, Cambridge University Press, 1991.

DEPARTMENT OF MATHEMATICS, BROWN UNIVERSITY, PROVIDENCE, RI 02912, USA
E-mail address: jglai@math.brown.edu

DEPARTMENT OF MATHEMATICS, BROWN UNIVERSITY, PROVIDENCE, RI 02912, USA
E-mail address: treil@math.brown.edu